



Redundant Array of Independent Disks (RAID)

RAID is an acronym for Redundant Array of Independent Disks. The term was coined in 1988 in a paper describing array configuration and application by researchers and authors Patterson, Gibson and Katz of the University of California at Berkeley.

In the past computer systems were often restricted to writing information to a single disk. This disk was usually expensive and prone to failure. Hard disks have always been the weakest link in computer systems, because the devices are the only mechanical member of an otherwise all- electronic system. The disk drive contains a mass of moving, mechanical parts operating at high speed. The question is not *whether* the hard drive will fail, but *when* a hard drive will fail.

RAID was designed to revolutionize the way computers managed and accessed mass storage of data by providing an inexpensive and redundant system of disks. It was called Redundant Array of Inexpensive Disks (RAID). Instead of writing to one Single Large Expensive Disk (SLED), RAID wrote to multiple inexpensive disks. Originally the name stood for Redundant Array of *Inexpensive* Disks but has been revised to *Independent* disks.

RAID Fundamentals

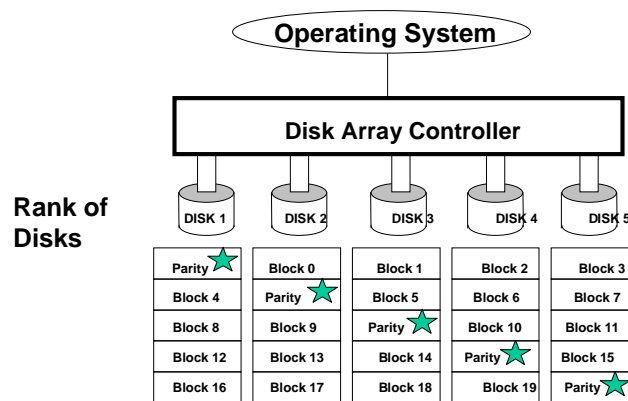
RAID accomplishes its goals of redundancy and fault tolerance by doing two things: one is striping and the other is parity checking. Striping means that files are written a block at a time over multiple disks. The striping technique divides data across many drives and improves data transfer rates and total disk transaction times. Such systems are good for transaction processing, but suffer from poor reliability because the system is only as reliable as the weakest individual drive.

Redundant Array of Independent Disks (RAID)

Parity checking ensures that the data is valid by performing a redundancy check on all data following a transmission. With parity, one of the disks on a RAID system can fail and the other disks have the ability to rebuild the failed disk. In both cases, these functions are transparent to the operating system. The Disk Array Controller (DAC) handles both striping and parity control.

Components

The major components in RAID are the Disk Array Controller (DAC) and a rank of five disks. The picture below shows an example of RAID-5. Data is striped across all five disks and the parity is used to recover a failed disk. There are many different RAID levels. Some RAID levels are designed for speed, some for protection, and some, like RAID-5 provide a combination of both. We will discuss them all.



Striping the Data

In the past a computer would write a file to a single disk. Striping allows you to break up a file and write different pieces to multiple disks concurrently. If you have 5 blocks of data in a file and stripe them across 5 disks, each block would be written to a separate disk simultaneously. If you had 5 OLTP transactions, each containing less than one block, you could process 5 different transactions concurrently.

Most RAID levels stripe at the block level but RAID can stripe at the bit or byte level. The size of the block is determined by the system administrator and is referred to as the stripe depth.

R e d u n d a n t A r r a y o f I n d e p e n d e n t D i s k s (R A I D)

To maximize a disk array subsystem's transaction processing capabilities, data must be written and read concurrently to and from multiple drives. To accomplish this, blocks of user data are striped across the array of drives. A stripe consists of a row of sectors (a sector consists of 512 bytes) located in the same position on each disk across the width of the array. Stripe depth, or the number of sectors in each data block, is defined by the subsystem software.

Stripe depth directly affects performance in that a too-shallow depth requires the system to execute more I/O commands than are needed. If the specified depth is too large, the processor's multi-tasking capabilities and the advantages provided by multiple drives and actuators may be negated.

In an ideal transaction environment, each request from the host involves only one drive, allowing multiple concurrent transactions to multiple drives.

The process of striping data to the array drives resolves the problem described earlier of overloading one system drive while another sits idle. Data striping eliminates the use of dedicated drives and ensures that the data processing load is balanced among the available drives, while increasing performance by writing multiple blocks concurrently.

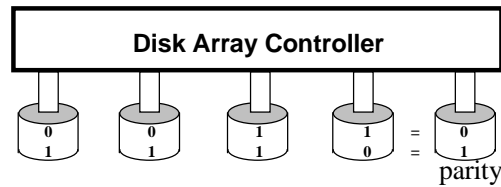
Parity

People often confuse parity with mirroring (or shadowing). Mirroring involves the creation of a duplicate copy of a disk. Mirroring is a technique where the data is written simultaneously to a pair of drives. These systems offer excellent reliability and have very good transaction processing results because the work can be carried out by either drive in the pair. The penalty paid is that two drives must be purchased to get the capacity of only one. The overhead of mirroring is 100% or double the disk space. When a disk fails the mirrored disk is used in its place.

Parity provides the same general protection as mirroring, but has less overhead. If a user has an array of five disks, four are used for data and one is used for parity. The overhead is only 20%. This is quite an advantage when cost is a concern.

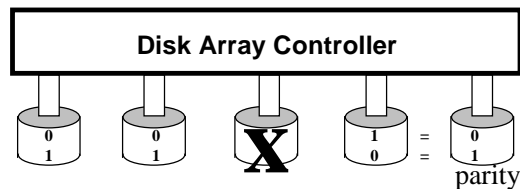
A computer writes only zeros or ones to represent data. A method to generate parity is called eXclusive OR (XOR). A bit is taken (either a 0 or 1) from each disk and totaled. If the total is even the parity bit is set to 0. If the total is odd the parity bit is set to 1. The picture below shows an example of the parity bits. The first four bits on the top of each disk add up to a total of 2. This even number causes the parity bit to be zero. The second four bits on the bottom of each disk add up to a total of 3. This odd number causes the parity bit to be one.

R e d u n d a n t A r r a y o f I n d e p e n d e n t D i s k s (R A I D)



Depending on the RAID level the parity will either be on one disk or be spread among all the disks. Either way it is 1/5th or 20% of the space when you are utilizing five disks. Parity is 1/4th or 25% of the space when utilizing four disks, and 1/3rd or 33% when utilizing three disks.

The picture below shows a RAID disk failure. Once the disk is replaced the Disk Array Controller will rebuild the disk to its previous contents. It will rebuild a 1 and a 1.



RAID Configuration Levels

The industry currently has agreed upon six RAID configuration levels and designated them as RAID 0 through RAID 5. Each RAID level is designed for speed, data protection, or a combination of both. The RAID levels are:

- **RAID - 0** Data striping Array
- **RAID - 1** Mirrored Disk Array
- **RAID - 2** Parallel Array, Hamming Code
- **RAID - 3** Parallel Array with Parity
- **RAID - 4** Independent Actuators with a dedicated Parity Drive
- **RAID - 5** Independent Actuators with parity spread across all drives

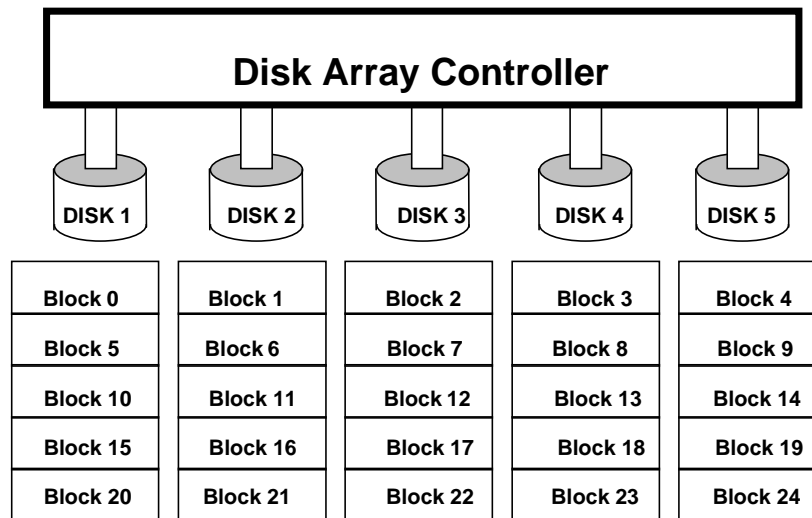
R e d u n d a n t A r r a y o f I n d e p e n d e n t D i s k s (R A I D)

The most popular RAID levels are RAID-0, RAID-1, and RAID-5. These are described next in more detail.

RAID-0 Data Striping Array

RAID-0 stripes the data across all the drives, but doesn't utilize parity. If one of the disks fails, the data must be restored on all five drives from backups. This RAID is designed for speed and is the fastest of all the RAIDs, but provides the least protection.

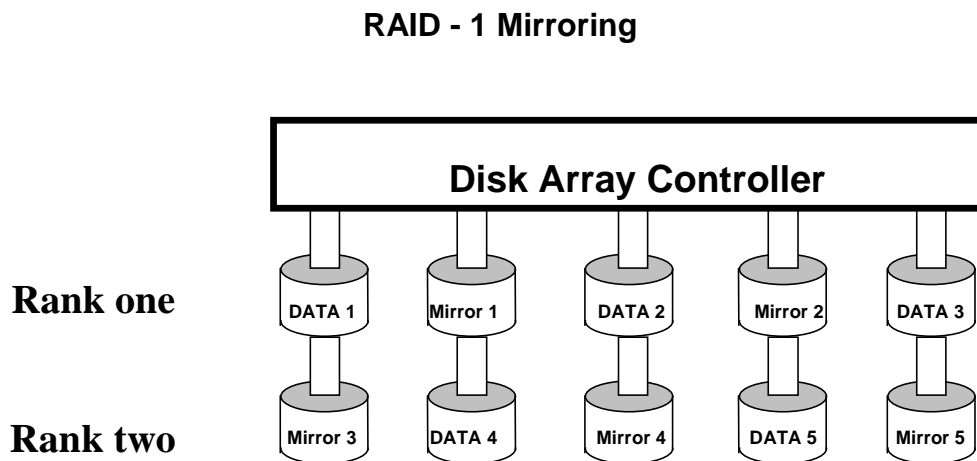
RAID - 0 Striping With No Parity



RAID - 1 Transparent or Striped Mirroring

The RAID-1 technology requires that each primary data disk have a mirrored disk. The contents of the primary disk and the mirror disk are identical. RAID-1 provides for the best data protection, but is slower than RAID 0 or 5.

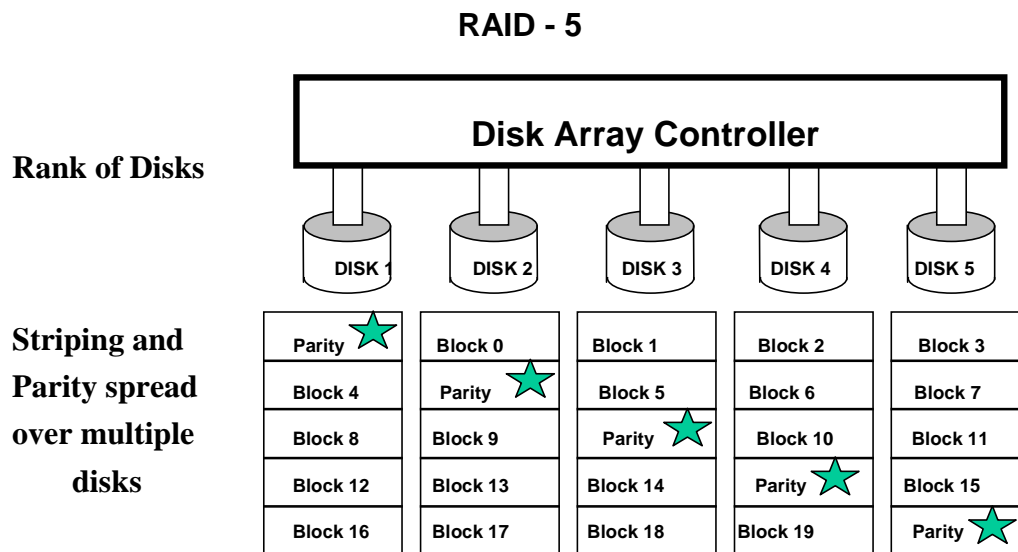
When data is written on the primary disk, a write also occurs on the mirror disk. The mirroring process is invisible to the user. For this reason, RAID-1 is also called transparent mirroring. The user can set up RAID-1 to write to a single disk and have that disk mirrored or can stripe to a number of disks with each of the striped disks also having a mirrored copy. This is called striped mirroring, RAID 1+0, RAID10, or in some cases RAID 6.



RAID-1 can utilize two, four, six, eight, or 10 disks

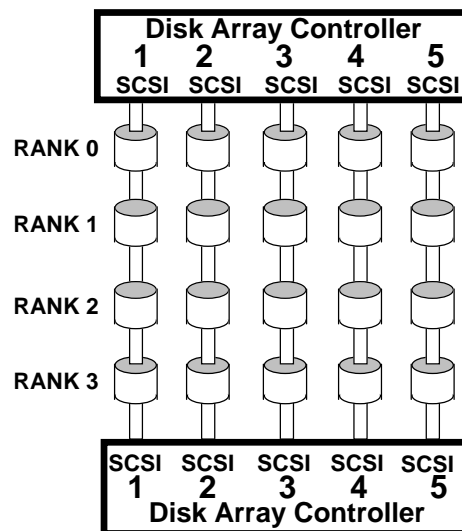
RAID - 5 Independent Actuators, Parity Spread

RAID-5 stripes data at the block level and also utilizes parity. With the RAID-5 technology, user information and parity are combined on every disk in the array. Independent and/or parallel data read and write operations are performed. This RAID is the most popular of all RAIDs. RAID-5 is not as fast as RAID-0 and does not provide as much protection as RAID-1 mirroring. RAID-5 provides good speed and good protection. This is why it is often the RAID level of choice.



RAID Disk Array Components

The major components of RAID Disk Arrays are the Disk Array Controllers, 5 SCSI Channels, and one or more ranks of disks. There are usually two Disk Array Controllers (DACs) working as a team. The implementation used to consist of one Active DAC and one Passive DAC in case of a DAC failure. Most implementations today utilize two Active DACs. They both share responsibilities for controlling the disks, but either one will control all ranks if the other DAC should fail. In the picture below you see two DACs. They share responsibility for four ranks. You can configure the disks with any supported RAID level. You can even break up the disks to configure multiple RAIDs within the same rank.



Internal/External Disk Arrays

In the past Disk Arrays were exclusively connected to the main computer through a cable and were always in an external box. There is a SCSI length limitation for external Disk Arrays of around 80ft or 25 meters. A repeater could be used for an additional 25 meters but that would result in a five-percent loss in performance.

Many computers today have internal RAID. The CPUs communicate with the disks internally, but the same fundamentals still exist. Whether internal or external, the disk array will have ranks of disks that are controlled by one or two disk array controllers.

Key Points to Remember

- RAID was created to enhance data performance, reliability and availability.
- Striping, parity checking and mirroring are three primary functions of RAID systems.
- RAID performs its functions transparent to the operating system.
- Systems are typically defined by ranks consisting of five disks each connected to one or two Disk Array Controllers.
- Different RAID levels provide varying degrees of speed and data protection.